Historic, Archive Document

Do not assume content reflects current
scientific knowledge, policies, or practices.

Data Services Center

Staff Paper

THE USE OF SPSS
IN THE
ECONOMIC DEVELOPMENT DIVISION
OF THE
U.S. DEPARTMENT OF AGRICULTURE

by

LINDA TOMPKINS and CLEVIE E. GLADNEY

THE USE OF SPSS
IN THE
ECONOMIC DEVELOPMENT DIVISION
OF THE
U.S. DEPARTMENT OF AGRICULTURE

by

LINDA TOMPKINS and CLEVIE E. GLADNEY

ESS Staff Report No. AGESS810528

THE USE OF SPSS IN THE ECONOMIC DEVELOPMENT DIVISION OF THE U.S.
DEPARTMENT OF AGRICULTURE. By Linda Tompkins and Clevie E. Gladney;
Data Services Center; Economics and Statistics Service; U.S. Depart-
ment of Agriculture; Washington, D.C. 20250; June 1981.

ABSTRACT

The paper examines the ways in which a commonly used statistical pack-
age, Statistical Package for the Social Sciences (SPSS), is used in
conducting research in such areas as indicators of rural well-being,
markets for rural labor, and the adequacy of community services (such
as water, sewerage, health, education, housing, energy, and transporta-
tion). Among the procedures and features to be included are CROSSTABS,
BREAKDOWN, REGRESSION, REPORT, data transformation techniques, and the
use of systems files.

Keywords: Economic and social research; Economic Development Division
(EDD); Statistical Package for the Social Sciences (SPSS).

```
* * * * * * * * * * * * * * * * * * * * * * * * * * *
* This paper was prepared for limited distribu-   *
* tion to the research community outside the      *
* U.S. Department of Agriculture.                 *
* * * * * * * * * * * * * * * * * * * * * * * * * * *
```

ACKNOWLEDGEMENTS

CONTENTS

I. OVERVIEW

The statistical analysis of agricultural and economic data, both
cross sectional and time series, has always been an important function
of the U.S. Department of Agriculture (USDA). In the past, researchers
could be found huddled over desk calculators, laboriously computing such
measures as correlation and regression coefficients. However, with the
advent of computers and statistical packages and programs, most of these
same researchers can now be found keying simple commands into terminals,
obtaining in seconds the identical results that once took hours, or even
days. This paper will examine the ways in which one of the most popular
of these statistical packages, Statistical Package for the Social Sciences
(SPSS), is used in conducting studies by a particular sector of researchers
at the USDA.

The Economic Development Division (EDD), a small unit of economists,
sociologists, and demographers within the economics research divisions
of the U.S. Department of Agriculture, conducts a national program of
economic and social research. This research "provides information to
improve the effectiveness of decisions of USDA and of other public agencies
that plan and manage the development process for different types of rural
communities. Specific research areas include demographic trends, markets
for rural labor, indicators of rural well-being, revenue sources for local
governments, adequacy of community services (such as water, sewerage, health,
education, housing, energy, credit, and transportation) and impact of Federal
programs and policies on rural growth."[1]

---

[1]ESC--Program Results and Plans, December 1979 (ESCS-76).

The use of SPSS by EDD began in the early seventies when the division purchased the package and installed it on the National Bureau of Standards' UNIVAC 1108. This purchase and subsequent use of SPSS eased the burden on the division's programming staff, allowed the researcher to obtain results in a shortened period of time, and provided a standard, documented, and reliable source of statistics for publications. When the division's files were later converted for use on the Washington Computer Center's (WCC) IBM 370, those researchers who previously used SPSS continued to do so, even though many other packages became available as a result of the conversion. Many of the researchers new to the division have a working knowledge of SPSS, having used the package on the various systems while employed elsewhere or through courses taught at many universities.

The Data Services Center (DSC), which provides computer and analytical support for the Agency, uses SPSS to maintain documentation for files (approximately 25 to date) received from various sources for use by EDD. The LIST FILEINFO and LIST ARCHINFO commands are used to provide researchers with a listing of variables names, variable and value labels, and missing codes for files created for their subsequent use. The largest county file currently maintained is the Bureau of Economic Analysis' Earnings and Employment files, which contains 10 years of data on 250 variables for approximately 3140 counties. Occasionally, FREQUENCIES, CONDESCRIPTIVE, and BREAKDOWN are used to perform preliminary data-checking and to determine the distribution of the variables to be analyzed.

II. SURVEY OF SELECTED EDD RESEARCHERS

EDD was chosen for this study because it was found that, among the economic research divisions at USDA, it contained the largest number of SPSS users. This is probably due to the fact that most of the socio-economic , demographic data collected and analyzed in the division is cross-sectional, while the other research divisions have the majority of their data in time series form. Not only that, but most of EDD's data are drawn from survey research and administrative records. We interviewed approximately twenty researchers and received the following information:

1. SPSS is mainly used because it is easy to learn, has excellent documentation, and contains those procedures and features best suited for the data being analyzed.

2. The most heavily used procedures are CROSSTABS, FREQUENCIES, BREAKDOWN, CONDESCRIPTIVE, REGRESSION, and REPORT. Features include RECODE, COMPUTE, SELECT IF, VAR and VALUE LABELS, and the creation of system and archive files.

3. Desired statistical techniques, not currently available in SPSS Release 8, are Cluster Analysis, Maximum Likelihood Estimation, Time Series Analysis, Multidimensional Scaling, and the capability to use the output from one procedure as the input to another, in the same job step.

4. Many of the national level data files analyzed use counties as the statistical case. There are approximately 3100 observations in these county files. Sources of data include

the Bureau of the Census, the Bureau of Economic Analysis,
the Bureau of Labor Statistics, and private vendors, such
as Dun and Bradstreet.

In general, SPSS is used in EDD to produce descriptive and inferential
statistics, such as the mean, standard deviation, median, correlation co-
efficients, percentages, and percentage changes. Because the results of
the various studies are too numerous and detailed to present in this paper,
the following selection of papers were abstracted to provide an overview
of the ways the output obtained from SPSS is used. Copies of any com-
plete study can be obtained by contacting the appropriate author.

THE RECENT SHIFT OF UNITED STATES POPULATION
TO NONMETROPOLITAN AREAS, 1970-75
By Calvin L. Beale

The purpose of this paper is to document the extent of population change from 1970-75 and to review the relative involvement of different types of counties in the trend. The REGRESSION subprogram in SPSS was used to perform a regression analysis of population change for 1970-75 and 1960-70 and other selected variables in nonmetropolitan counties. This procedure showed the variables to have different degrees of association with population change by region. Revived population growth in rural and small town areas, approximated by the use of nonmetropolitan counties, was shown to have continued through 1975.

PROBLEMS OF RURAL ELDERLY HOUSEHOLDS IN POWELL COUNTY, KENTUCKY
By Donald K. Larson and E. Grant Youmans

This study shows that many elderly households in Powell County had incomes below the national poverty level, with Social Security being the most frequently reported source of income. In addition to income, the other basic problem areas cited were the need for home repair and transportation. Various crosstabulations were constructed with SPSS involving household types, income classes, sources of income, type and size of residences, housing needs, and modes of transportation in order to compute statistics and make inferences based on the chi-square test.

THE HIRED FARM WORKING FORCE OF 1977

By Gene Rowe

This report presents data on the demographic, social, and economic characteristics of persons 14 years of age and over, who did farmwork for cash wages or salaries during 1977. The FREQUENCIES procedure was used to determine the median values for age, number of years employed in farmwork, and educational level of hired farmworkers. It was found that the median age was 30 years for Hispanics and 33 years for Blacks and others, compared to 23 years for Whites. These results suggest that white workers are more likely than others to move out of hired farmwork as they become older and use this work only as an entrance into the labor force.

INDICATORS OF SOCIAL WELL-BEING FOR U.S. COUNTIES

By Peggy J. Ross, Herman Bluestone, and Fred K. Hines

This study constructs four composite indexes to measure or monitor the many dimensions of social well-being: socioeconomic, health status, family status, and alienation. The report then uses these indexes to measure intrastate and regional variations among metropolitan and non-metropolitan counties. The FACTOR and BREAKDOWN subprograms in SPSS were used to derive these scores through principal component analysis and to compute mean index scores of counties grouped by metro-nonmetro status and rural urban orientation respectively.

STRUCTURAL DIFFERENTIATION AND RURAL DEVELOPMENT

By Richard G. Stuby

.

This study examines the concept of structural differentiation, used to measure differences between the levels of development in rural U.S. counties. The SPSS procedure GUTTMAN is used to compute the Guttman scale of structural differentiation, derived from Dun and Bradstreet retail establishment data, which is then examined for its validity as a measure and its ability to discriminate among counties.

USE OF MULTIPLE REGRESSION ANALYSIS TO SUMMARIZE AND INTERPRET
LINEAR PROGRAMMING SHADOW PRICES IN AN ECONOMIC PLANNING MODEL

By Daniel G. Williams

A simple method is presented for evaluating the benefit to a region (regional objective function) of new manufacturing firms. These firms are subsets of the more aggregated 4-digit Standard Industrial Classification manufacturing industries in a rural multicounty economic planning model. SPSS performs the multiple regression analysis used to summarize and interpret shadow prices of export industries so that local planners in their decisionmaking can use the underlying economic characteristics of the model industries rather than use only their industry product classifications.

III. AIDS FOR SPSS USERS

This section will examine tools available to DSC and the research
staff in our Agency. These tools were prepared by our data processing
and consulting staff to disperse information about SPSS, and to facilitate
its use.

A. "HOW TO" PAPER SERIES

The "HOW TO" paper series provides computer users with short guides
on performing various data processing tasks. Each "HOW TO" paper is one
to two pages in length and covers a topic such as using a major statistical
package or programming language (e.g., SPSS, SAS, SPEAKEASY) or accessing
data from the Agency's data bank. The following papers involve using SPSS:

1. HOW TO...USE SPSS

2. HOW TO...SAVE AND RETRIEVE A SPSS DATA FILE

3. HOW TO...USE SPSSXEQT

The latter refers to a user-oriented command procedure which will be discussed
in Section III C.

HOW TO ...

ESCS
DATA SERVICES CENTER
USDA

USE SPSS
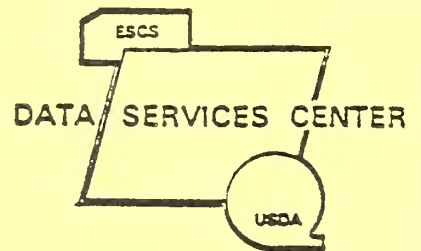
The Statistical Package for the Social Sciences (SPSS) is an integrated system of computer programs designed for the analysis of social science data which enables the user to perform various types of analyses in a simple manner. Since it is primarily used in a batch environment, users must prepare Job Control Language (JCL) according to the requirements of their programs.

* *

I. JCL FOR A SIMPLE JOB - DATA ARE ON CARDS

1. Write and keypunch your SPSS control statements and data on cards.
2. Keypunch the following job control cards, starting in column 1:

    a. //ERSbiijj JOB (XXXX,RJØ13),'programmer',CLASS=k,TIME=(mm,ss)

         b - bin number
         ii - programmer's initials
         jj - code for the job (determined by you)
         XXXX - 10 digit account number
         k - class code
         mm,ss - time allotted to the job in minutes and seconds

        Example: //ERS8MWSS JOB (XXXX,RJØ13),'SPSS JOB',CLASS=C,TIME=(,2Ø)

    b. //*LOGONID logon-id

    c. //*PASSWORD associated-password

    d. //stepname EXEC SPSSH

        "stepname" is any name chosen by you; for example, STEPØ1

    e. //SYSIN DD *

    f. //

    g. ||

* *

II. JCL FOR USING RAW DATA ON DISK OR TAPE

1. To use data stored on tape, punch the following cards:

    //FTØ8FØØ1   DD DSN=data.set.name,UNIT=TAPE,DISP=OLD,
    //   VOL=SER=serial-number

    "serial-number" is the reel number of the tape, which is usually given on the job log when the dataset was created on tape.

2. To use data stored on disk, punch the following cards:

```
//FTØ8FØØ1    DD  DSN=data.set.name,UNIT=SYSDA,DISP=SHR
```

Insert these cards after the EXEC SPSSH card as indicated in I above.

* *

III. <u>JCL FOR OUTPUTTING RAW DATA ONTO DISK OR TAPE</u>

1. To set up a raw data file on tape, punch the following cards:

```
//FTØ9FØØ1  DD  DSN=data.set.name,UNIT=TAPE,DISP=(NEW,KEEP),
//   DCB=(RECFM=FB,LRECL=xxx,BLKSIZE=yyy)
```

2. To set up a raw data file on disk, punch the following cards:

```
//FTØ9FØØ1  DD  DSN=data.set.name,UNIT=SYSDA,DISP=(NEW,CATLG),
//   DCB=(RECFM=FB,LRECL=xxx,BLKSIZE=yyy),SPACE=(yyy,(ww,ww),RLSE)
```

(where xxx , ww ,yyy are supplied by the user.)

Insert these cards after the EXEC SPSSH card. When output is returned, make note of the VOL=SER numbers so that these data files can be easily used again.

* *

IV. <u>JCL FOR SORTING DATA</u>

When using a SORT CASES control card in your SPSS program, you will need to punch additional JCL cards:

1. `//SORTLIB   DD DSN=SYS1.SORTLIB,DISP=SHR`
2. `//SYSOUT    DD SYSOUT=A`
3. `//SORTWKØ1  DD UNIT=SYSDA,SPACE=(CYL,(20,20))`
4. `//SORTWKØ2  DD UNIT=SYSDA,SPACE=(CYL,(20,20))`
5. `//SORTWKØ3  DD UNIT=SYSDA,SPACE=(CYL,(20,20))`

Insert these cards immediately before the "//SYSIN DD *" card.

* *

V. <u>FOR FURTHER INFORMATION</u>

1. Consult the SPSS manual, <u>Statistical Package for the Social Sciences,</u> for help with SPSS and JCL statements.

2. Refer to "HOW TO... SAVE AND RETRIEVE A SPSS DATA FILE" and "HOW TO USE SPSSXEQT."

3. Contact T.C. Lu or Linda Tompkins, Rm 103, X77577, of the Stat/Math Team.

HOW TO ...
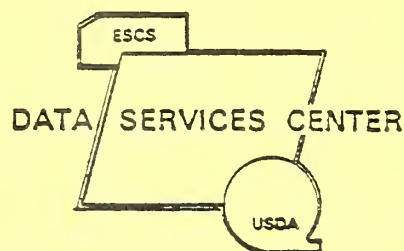
SAVE AND RETRIEVE A SPSS DATA FILE

Conversion of raw data files into files that can be easily accessed by SPSS (SPSS data files) is a process that allows SPSS programs to run more efficiently and lets users save computed results for future use. SPSS data files can also be read and used by SAS in PROC CONVERT and by Speakeasy with the GETSPSS command. Establishing SPSS data files is, however, only recommended if the same data is to be used more than once.

* *

I. STORING A SPSS DATA FILE

In order to store raw data in a SPSS data file, your SPSS program must contain both a FILE NAME card and a SAVE FILE card (if more than 500 variables are to be stored, the SAVE FILE card is replaced by a SAVE ARCHIVE card).

> FILE NAME   XXXXXXX
>   .
>   .
> SAVE FILE

The SAVE FILE card is placed immediately before the FINISH card.

The JCL necessary to run a SPSS program should be set up as instructed in "HOW TO. . .USE SPSS" and in the SPSS user's manual, with the additional statement:

1. To set up a file on tape, punch the following cards:

   ```
   //FT04F001  DD UNIT=TAPE,DSN=user.name,DISP=(NEW,KEEP),
   // DCB=BLKSIZE=2012
   ```

   NOTE: In order to use this file again, make note of the VOL=SER number on the JCL listing of your run. To extend the life of your file beyond 5 days, execute the KPTAPES procedure.

2. To set up a file on disk, punch the following cards:

```
//FTØ4FØØ1  DD UNIT=SYSDA,DSN=user.name,DCB=BLKSIZE=2012,
//  DISP=(NEW,CATLG,DELETE),SPACE=(2012,(1,),RLSE)
```

These cards should follow the EXEC SPSS card and should precede the FTØ5FØØ1 input card. The file will be stored in the data set labeled by 'user.name' (a conventional data set name; for example, ERPRE.ERBIN.FILENAME.DATA).

* *

II. RETRIEVING A SPSS DATA FILE

In order to retrieve a SPSS data file, your SPSS program must contain a GET FILE card of the form

GET FILE   XXXXXXXX

where XXXXXXXX is the file name used when the file was saved. This card will be the first of your SPSS program cards unless a RUN NAME card is used. (In cases where you need to retrieve variables from an archive file, GET ARCHIVE replaces GET FILE.)

JCL necessary to run SPSS must be set up, including the following statement:

1. To retrieve a tape data file, punch the following cards:

```
//FTØ3FØØ1 DD UNIT=TAPE,DSN=user.name,VOL=SER=YYYY,
//  DISP=(OLD,KEEP)
```

where YYYY is the VOL=SER number listed when the data file was created.

2. To retrieve a disk data file, punch the following cards:
```
//FTØ3FØØ1 DD DSN=user.name,DISP=SHR,UNIT=SYSDA,
//  VOL=SER=STORXX
```
Where STORXX can be obtained from the run creating the SPSS data file, e.g VOL=SER=STORL5.
These cards should follow the EXEC SPSS card and must precede the FTØ5FØØ1 input card. 'User.name' is the name used when the data was stored.
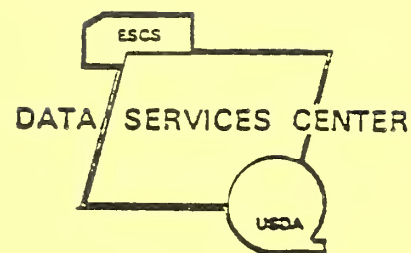
III. ORDER OF JCL CARDS
It is important that the JCL cards for data sets used in an SPSS run appear in "ascending" order, i.e. FTØ3FØØ1 precedes FTØ4FØØ1, etc.

IV. FOR FURTHER INFORMATION
1. Reference the SPSS users manual, p.78-88 and p.585-603.
2. Consult "HOW TO ... USE SPSS" and 'HOW TO ... USE SPSSXEQT"
3. Contact Linda Tompkins, Rm. 103, X77577.

HOW TO ...

USE SPSSXEQT

SPSSXEQT is a procedure available on TSO that will submit a batch SPSS job. You must have the following information at hand before executing the procedure:

1. The complete data set name(s) of your input data file(s) and the associated tape number(s) if the data is stored on tape(s).

2. If you will be creating an output file on disk, you must know the approximate number of cases to be written and the number of variables to be saved.

3. The complete data set name of the file containing your SPSS control statements.

4. A jobname of the form ERSniisj

5. Your WCC account number

To invoke the procedure:

1. Logon to TSO, wait for READY.

2. Key in SPSSXEQT. This will initiate the procedure for you.

3. Answer each question asked. After you have responded to all questions, your job will be automatically submitted. If you have made a mistake in answering any questions, you will have an opportunity to cancel the job and start over.

4. Depending upon the option you select, your output may be picked up in the terminal room or may be retrieved at your terminal when it is ready using the OUTPUT command.

All question or suggestions for improvement of the CLIST should be directed to Clevie Gladney (447-8389) or Dick Stuby (447-8877).

For information on the SPSS package, interested persons should contact T.C. Lu or Linda Tompkins, Rm. 103, X77577.

B. SPSS WORKSHOPS

We determined from our interviews that most researchers learned SPSS
through self-instruction. DSC does sponsor workshops, approximately
every four months, which introduce the participants to the fundamentals
of computing with SPSS. These workshops consist of three two-hour
sessions, attended by fifteen to twenty researchers and staff and contain
those features and procedures which were found to be most useful to the
researchers in the Agency (e.g., CROSSTABS and BREAKDOWN are used
extensively, whereas SURVIVAL and RELIABILITY are not).

The workshops have proven to be successful, since most participants
have been able to subsequently produce sophisticated SPSS programs with
minimal assistance. The following pages contain the workshop syllabus
and homework assignments. We also make a special effort to distinguish
among our available statistical packages (which include SPSS, SAS, BMDP,
TSP, and SPEAKEASY) in regard to such issues as potential and actual
research applications, statistical features, ease of use, and accessibility.
Other handouts that exhibit examples of successfully executed procedures
and features are also given at the workshop, but are too lengthy to include
here.

<u>S.P.S.S. WORKSHOP</u>

Session I

I. General Description of S.P.S.S.

    a. Features
    b. Advantages and Disadvantages
    c. Control Card Preparation

II. Run Cards

    a. RUN NAME
    b. TASK NAME
    c. COMMENT
    d. FINISH
    e. Error Messages - EDIT

III. Data-Definition Cards - Needed to describe the input data set to the system.

    a. FILE NAME
    b. VARIABLE LIST
    c. INPUT FORMAT
    d. DATA LIST
    e. INPUT MEDIUM
    f. N OF CASES
    g. READ INPUT DATA and END INPUT DATA

IV. A Few Task-Definition Cards - Activate, define and control the calculations
to be performed on the data

    a. FREQUENCIES
    b. OPTIONS and STATISTICS
    c. CROSSTABS

V. HOMEWORK

    a. J.C.L. Statements
    b. Clist Procedures

I. Data-Definition Cards (cont.)

    a. VAR LABELS
    b. VALUE LABELS
    c. MISSING VALUES
    d. ASSIGN MISSING

II. Data Modification and Selection Cards - Allow the user to create new or modify existing variables

    a. RECODE
    b. COMPUTE
    c. LAG
    d. IF
    e. SELECT IF
    f. Temporary vs. Permanent Modification

III. More Task-Definition Cards

    a. CONDESCRIPTIVE
    b. BREAKDOWN
    c. REGRESSION

IV. HOMEWORK

Session III

I. Other Statements and Procedures

    a. DO REPEAT - END REPEAT
    b. SUBFILE LIST - RUN SUBFILES
    c. WRITE CASES
    d. REPORT
    e. MULT RESPONSE

II. S.P.S.S System Files

    a. Creation - FILE NAME
                 SAVE FILE
    b. Retrieval - GET FILE
    c. Documentation - LIST FILEINFO
    d. Modification - DELETE VARS
                     KEEP VARS
                     ADD VARIABLES
                     ADD DATA LIST
                     ADD CASES
    e. Archive Files
    f. SAS and Speakeasy Conversion Routines

III. Conclusions

PROBLEM I

There exists a data set stored on disk which contains the following eight variables:
(25 observations)

Var. No.	Description	Location
1	Age (Actual number coded)	6-10
2	Sex 1 Male 2 Female	11-15
3	Race 1 White 2 Black 3 Other	16-20
4	Marital Status 1 Married 2 Widowed 3 Divorced or separated 4 Never married	21-25
5	Relation to Head of Household 1 Spouse 2 Child of Head 3 Parent of Head or Spouse 4 Other relative 5 Others (not related) 6 Head of Household	26-30
6	Income Level 1 $0 - 7,999 2 $8,000 - 15,999 3 $16,000 - 24,999 4 $25,000 or more	31-35
7	Amount in savings account (actual number coded) 999999=no response	36-43
8	Amount in checking account (actual number coded) 888888=no response	44-51

You are to:

1. Use the file: ERS73.ER8LT.HOMEWORK.DATA

2. Create frequency tables and histograms for variables 2, 3, and 4.

3. Create frequency tables for variables 5 and 6 and get the minimum and maximum values.

4. Construct crosstabulations that will show the following information:

    a.   the number of white females

    b.   the marital status of heads of households

    c.   the income category which contains the greatest number of white males

```
E
L
        23      1      1      4      2      1      200        50
        40      2      1      2      6      4   999999    888888
        19      1      2      3      2      1     5000       700
        30      1      3      1      1      3      700       100
        20      1      3      4      2      1        0         0
        21      2      2      2      4      1      400       200
        33      2      2      3      4      2      160        20
        38      1      1      1      1      4       50        20
        46      2      3      1      6      2   999999        40
        50      2      1      1      6      4    10000    888888
        32      2      1      1      6      4      300         0
        27      2      1      1      1      2       60       100
        20      1      2      4      2      2        0        50
        51      2      3      3      3      4     5700        80
        46      1      2      1      1      3      900    888888
        29      2      1      3      6      2       80        30
        52      2      2      2      5      4     3000    888888
        41      1      1      3      6      3      200       200
        19      1      1      2      2      1       25         0
        24      2      1      1      1      2     2000       100
        36      2      1      4      5      1      100        50
        47      1      3      3      3      3      200        75
        53      1      2      1      6      4     9800       200
        34      1      2      1      1      3      800        50
        43      1      1      3      3      3   999999    888888
END OF DATA
E
        DATA SET NAME - 'ERS73.ER8LT.HOMEWORK.DATA'
SUBCOMMAND DATA NOT FOUND
E
END SAVE
READY
```

PROBLEM II

Using the data set from Problem I, you are to:

1. Create a new variable as follows:

    Age categories
    1   under 21
    2   21-35
    3   over 35

2. Find the average and range for the amount of assets.   (Do not include
    'no response' values!).   Assets = Savings + Checkings amounts.

3. Find the average amount of money in savings and checking accounts
    for:

    a.   white females in the above age categories
    b.   married heads of households in the above age categories

4. Regress income levels on variables 1 and 2.   Also obtain the
    correlation matrix.   Then repeat the regression for age category
    3 only.

5. Label all variables and their values used above.

C. TSO COMMAND PROCEDURES FOR BATCH JOB SUBMISSION

Many researchers find IBM 370 Job Control Language (JCL) very difficult to comprehend. In creating procedures for batch job submission via Time Sharing Option (TSO), it was our objective to make the use of the computer as painless as possible for SPSS users. Several command procedures were written to minimize the problems frequently encountered in using JCL. The three current procedures are described below.

ISPSS is used to submit "simple" batch jobs. It requires an on-line file containing the SPSS statements and another on-line file containing the input data. The output from the run is routed to the remote job entry station for printing.

CSPSS interactively edits SPSS control statements for syntax errors by inserting an EDIT statement in front of the user's SPSS program and executing the job in real time. The results are printed back to the user's terminal. If there are errors, the user can correct them and execute CSPSS again. If not, the program can be run by executing either ISPSS or SPSSXEQT.

SPSSXEQT is an enhanced version of ISPSS. Whereas ISPSS has no output facilities and the restriction of disk data input, SPSSXEQT has the capability of accepting various inputs and creating all types of output usually associated with SPSS jobs.

Examples of runs using the three procedures follow, where the user's responses to prompts are underlined.

EXAMPLE 1

```
SPSS
WELCOME TO SPSS ERS34
THIS PROCEDURE EXECUTES SPSS INTERACTIVELY.  YOUR SPSS CONTROL
CARDS AND DATA MUST BE IN SEPARATE DISK FILES.
NOW ENTER YOUR SPSS CONTROL CARDS FILE FULL DATA SET NAME:
ers34.er7cg.spss.control.data
TYPE IN YOUR DATA FILE'S FULL DATASET NAME:
ers73.er91t.test.data
YOUR RUN IS FINISHED AND RESULTS, LABELED ERS34, ARE BEING
PRINTED DOWNSTAIRS IN THE RJE TERMINAL ROOM.
READY
```

EXAMPLE 2

```
SPSS
WELCOME TO DSC'S INTERACTIVE SPSS ERS34
THIS PROCEDURE 'EDITS' YOUR SPSS CONTROL CARDS FOR ERRORS.
NOW ENTER YOUR SPSS CONTROL CARDS FILE FULL DATA SET NAME:
ers34.er7cg.spss.control.data
SPSS BATCH SYSTEM                                                        10/01/80          PAGE

SPSS FOR OS/360, VERSION H, RELEASE 8.0, FEBRUARY 15, 1979
DEFAULT SPACE ALLOCATION..    ALLOWS FOR..    102 TRANSFORMATIONS
WORKSPACE      71680 BYTES                    409 RECODE VALUES + LAG VARIABLES
TRANSPACE      10240 BYTES                   1641 IF/COMPUTE OPERATIONS

            1 EDIT
            2 RUN NAME       SPSS - SPSSXEQT TEST RUN
            3 VARIABLE LIST  X1 TO X5
            4 INPUT MEDIUM   DISK
            5 INPUT FORMAT   FIXED  (F1.0,2F7.1,F6.1,2X,A1)


            ACCORDING TO YOUR INPUT FORMAT, VARIABLES ARE TO BE READ AS FOLLOWS

            VARIABLE   FORMAT   RECORD    COLUMNS

            X1         F 1. 0      1       1-    1
            X2         F 7. 1      1       2-    8
            X3         F 7. 1      1       9-   15
            X4         F 6. 1      1      16-   21
            X5         A 1         1      24-   24
THE INPUT FORMAT PROVIDES FOR   5 VARIABLES.   5 WILL BE READ
IT PROVIDES FOR  1 RECORDS ('CARDS') PER CASE.  A MAXIMUM OF   24 'COLUMNS' ARE USED ON A RECORD.

            6 VAR LABELS     X1    MARITAL STATUS/
            7                X2    BIWEEKLY SALARY/
            8                X3    MONTHLY RENT/
            9                X4    AVERAGE GROCERY BILL/
           10                X5    GRADE IN COURSE
           11 VALUE LABELS   X1   (1) MARRIED (2) WIDOWED
           12                     (3) DIVORCED (4) SEPARATED
           13                     (5) NEVER MARRIED
           14 CONDESCRIPTIVE  X2 TO X4

***** GIVEN WORKSPACE ALLOWS FOR 1433 VARIABLES FOR CONDESCRIPTIVE PROBLEM ****


           15 FINISH
      USAGE DATA FILE IS FOR WRONG VERSION OF SPSS
      NOTIFY YOUR SPSS COORDINATOR OF THIS ERROR.
      NORMAL END OF JOB.
         15 CONTROL CARDS WERE PROCESSED.
          0 ERRORS WERE DETECTED.
YOUR 'COMPILE' IS FINISHED ERS34
READY
```

EXAMPLE 3

```
.READY
 SPSSXEQT
 THIS PROCEDURE SUBMITS A BATCH SPSS RUN FOR YOU ERS188 WHEN ASKED A QUESTION,
 PLEASE RESPOND WITH THE APPROPRIATE CODE NUMBER OR KEYWORD.
 IF YOU ENTER A BAD CODE BY MISTAKE, HIT THE BREAK KEY,
 THEN THE RETURN KEY.   THIS WILL ALLOW YOU TO EXIT THE CLIST AND START OVER.
 ALL FILES OUTPUT TO TAPE WILL BE AUTOMATICALLY SET TO EXPIRE IN 60 DAYS
 FROM 09/17/80
 YOU MUST SUBMIT A KEEPTAPE IN ORDER TO EXTEND THE EXPIRATION DATE.
 FIRST WE NEED SOME INFORMATION ON THE DATA TO BE USED IN THIS JOB.
 ENTER THE NUMBER THAT IDENTIFIES YOUR INPUT DATA:
     1 RAW DATA FILE (REQUIRES VARIABLE LIST AND INPUT FORMAT)
     2 SPSS SYSTEM FILE (REQUIRES GET FILE CARD)
     3 SPSS ARCHIVE (REQUIRES GET ARCHIVE CARD)
     4 TWO SPSS SYSTEM FILES, TWO ARCHIVE FILES, OR ONE SPSS SYSTEM FILE
       AND ONE ARCHIVE FILE.
     5 A SYSTEM FILE AND A RAW INPUT FILE OR AN ARCHIVE AND A RAW INPUT FILE.
 ENTER NUMBER.
 4
 AT THE PRESENT TIME YOU MAY CALL FOR TWO SPSS SYSTEM FILES,
 TWO ARCHIVE FILES OR ONE ARCHIVE AND ONE SYSTEM FILE.   NOW ENTER THE COMPLETE
 DATA SET NAME OF EACH OF YOUR INPUT FILES.   (ONE PER LINE)
 ERS99.ER7IN.SYSTEM.FILE1.DATA
 ERS99.ER7IN.SYSTEM.FILE2.DATA
 WHERE ARE BOTH OF YOUR FILES STORED?  ENTER A CODE FROM BELOW:
     1   BOTH FILES ON TAPE
     2   BOTH FILES ON DISK
     3   ONE ON TAPE ONE ON DISK
 3
 IS ERS99.ER7IN.SYSTEM.FILE1.DATA STORED ON TAPE OR DISK?
 ENTER EITHER TAPE OR DISK.
 DISK
 ENTER THE 6 DIGIT REEL NUMBER OR NUMBERS SEPARATED BY A COMMA AND ENCLOSED
 IN PARENTHESES FOR ERS99.ER7IN.SYSTEM.FILE2.DATA
 IE. (123456) OR (123456,098765,345678,...)
 (468291)
 DO YOU WISH TO SAVE AN OUTPUT FILE AT THE END OF THIS RUN?
 ENTER YES OR NO.
 YES
 ENTER SOME INFORMATION FOR THE TYPE OF OUTPUT FILES
 THAT YOU WILL BE GENERATING IN YOUR SPSS JOB.
 SELECT THE NUMBER THAT IDENTIFIES THE TYPE OF OUTPUT FILE YOU WANT.
     1 AN SPSS SYSTEM FILE (REQUIRES A SAVE FILE CARD) (CHOICE OF TAPE OR DISK)
     2 AN SCSS (CONVERSATIONAL SYSTEM) MASTER FILE (REQUIRES A SAVE SCSS CARD)
                 (FILES TO DISK ONLY)
     3 RAW DATA FILE (REQUIRES A WRITE CASES OR AGGREGATE CARD)(ALL FILES TO TAPE)
 ENTER NUMBER
 1
 DO YOU WANT YOUR SYSTEM FILE TO BE STORED ON TAPE OR DISK?
 ENTER TAPE OR DISK.
 DISK
 ENTER THE COMPLETE DATA SET NAME FOR YOUR OUTPUT FILE  BE SURE TO BEGIN WITH
 A VALID LOGONID FOLLOWED BY YOUR JOBID: EG. ERS99.ER7ZZ..... ETC.
 ERS99.ER7OT.SYTEM.MERGED.FILE.DATA
 HOW MANY VARIABLES WILL YOU BE SAVING ON YOUR OUTPUT FILE?
 125
 HOW MANY CASES WILL BE ON YOUR OUTPUT FILE?
 3144
 YOU EITHER HAVE TOO MANY VARIABLES OR TOO MANY CASES.
 AS IT STANDS NOW, YOU ARE REQUESTING 167 TRACKS AND A MAXIMUM OF 95 IS ALLOWED.
 YOU MAY NOT USE THIS CLIST. HIT THE BREAK KEY THEN THE RETURN TO GET OUT
 SEE CLEVIE GLADNEY (7-8389) OR DICK STUBY (7-8877) FOR INFORMATION ON HOW TO
 SOLVE THIS PROBLEM.
```

EXAMPLE 4

```
READY
SPSSXEQT
THIS PROCEDURE SUBMITS A BATCH SPSS RUN FOR YOU ERS188 WHEN ASKED A QUESTION,
PLEASE RESPOND WITH THE APPROPRIATE CODE NUMBER OR KEYWORD.
IF YOU ENTER A BAD CODE BY MISTAKE, HIT THE BREAK KEY,
THEN THE RETURN KEY.   THIS WILL ALLOW YOU TO EXIT THE CLIST AND START OVER.
ALL FILES OUTPUT TO TAPE WILL BE AUTOMATICALLY SET TO EXPIRE IN 60 DAYS
FROM 09/17/80
YOU MUST SUBMIT A KEEPTAPE IN ORDER TO EXTEND THE EXPIRATION DATE.
FIRST WE NEED SOME INFORMATION ON THE DATA TO BE USED IN THIS JOB.
ENTER THE NUMBER THAT IDENTIFIES YOUR INPUT DATA:
    1 RAW DATA FILE (REQUIRES VARIABLE LIST AND INPUT FORMAT)
    2 SPSS SYSTEM FILE (REQUIRES GET FILE CARD)
    3 SPSS ARCHIVE (REQUIRES GET ARCHIVE CARD)
    4 TWO SPSS SYSTEM FILES, TWO ARCHIVE FILES, OR ONE SPSS SYSTEM FILE
      AND ONE ARCHIVE FILE.
    5 A SYSTEM FILE AND A RAW INPUT FILE OR AN ARCHIVE AND A RAW INPUT FILE.
ENTER NUMBER.
4
AT THE PRESENT TIME YOU MAY CALL FOR TWO SPSS SYSTEM FILES,
TWO ARCHIVE FILES OR ONE ARCHIVE AND ONE SYSTEM FILE.   NOW ENTER THE COMPLETE
DATA SET NAME OF EACH OF YOUR INPUT FILES.   (ONE PER LINE)
ERS99.ER7IN.SYSTEM.FILE1.DATA
ERS99.ER7IN.SYSTEM.FILE2.DATA
WHERE ARE BOTH OF YOUR FILES STORED?   ENTER A CODE FROM BELOW:
    1   BOTH FILES ON TAPE
    2   BOTH FILES ON DISK
    3   ONE ON TAPE ONE ON DISK
2
DO YOU WISH TO SAVE AN OUTPUT FILE AT THE END OF THIS RUN?
ENTER YES OR NO.
YES
ENTER SOME INFORMATION FOR THE TYPE OF OUTPUT FILES
THAT YOU WILL BE GENERATING IN YOUR SPSS JOB.
SELECT THE NUMBER THAT IDENTIFIES THE TYPE OF OUTPUT FILE YOU WANT.
    1 AN SPSS SYSTEM FILE (REQUIRES A SAVE FILE CARD) (CHOICE OF TAPE OR DISK)
    2 AN SCSS (CONVERSATIONAL SYSTEM) MASTER FILE (REQUIRES A SAVE SCSS CARD)
              (FILES TO DISK ONLY)
    3 RAW DATA FILE (REQUIRES A WRITE CASES OR AGGREGATE CARD)(ALL FILES TO TAPE)
ENTER NUMBER
1
DO YOU WANT YOUR SYSTEM FILE TO BE STORED ON TAPE OR DISK?
ENTER TAPE OR DISK.
DISK
ENTER THE COMPLETE DATA SET NAME FOR YOUR OUTPUT FILE  BE SURE TO BEGIN WITH
A VALID LOGONID FOLLOWED BY YOUR JOBID: EG. ERS99.ER7ZZ..... ETC.
ERS99.ER7OT.MERGED.SYSTEM.FILE.DATA
HOW MANY VARIABLES WILL YOU BE SAVING ON YOUR OUTPUT FILE?
45
  HOW MANY CASES WILL BE ON YOUR OUTPUT FILE?
600
```

EXAMPLE 4 (cont.)

```
NOW YOU ARE READY TO SUBMIT YOUR JOB.
ENTER THE DATA SET NAME FOR YOUR SPSS CONTROL STATEMENTS.
ERS99.ER7XX.MERGE.FILES.CONTROL.DATA
NOW ENTER A JOBNAME TO IDENTIFY YOUR RUN.  NOTE:  JOB NAMES HAVE UP TO 8
CHARACTERS AND MUST BEGIN WITH ERS; FOR EXAMPLE ERS7XXXX.
ERS7XX01
ENTER YOUR 10 DIGIT ACCOUNT NUMBER IE 415001XXXX
4150010501
NOW ENTER YOUR JOBID (THE SECOND LEVEL QUALIFIER USED IN YOUR DATA SET NAME).
FOR EXAMPLE:  ER7XX
ER7XX
DO YOU WANT YOUR JOB ROUTED TO
  1  LINE PRINTER (PICKUP OUTPUT IN ROOM 192)
  2  HELD ON OUTPUT QUEUE (SELECT THIS OPTION ONLY IF YOU KNOW HOW
     TO RELEASE THE QUEUE)
ENTER NUMBER
1
ENTER ONE OF THE FOLLOWING JOB PRIORITY CODES:
        1 = OVERNIGHT (PREFERRED PRIORITY CATEGORY  --  SAVES BUCKS)
        3 = NORMAL
       13 = HIGH (CONSULT WITH DATA COORDINATOR BEFORE USE)
1
 ENTER A JOB CLASS BASED ON THE FOLLOWING:
CLASS B FOR 20 SECONDS OF CPU TIME  AND 3 OR LESS TAPES
CLASS I FOR 2 MINUTES OF CPU TIME  AND 6 OR LESS TAPES
CLASS N FOR 10 MINUTES OF CPU TIME AND 6 OR LESS TAPES
CLASS H FOR 10 MINUTES  AND 16 OR LESS TAPES
CLASS C FOR 20 SECONDS OF CPU TIME AND NO TAPES
CLASS D FOR 2 MINUTES  AND  NO TAPES
 ENTER ONE OF THE SIX CLASSES (B I N H C D ).
D
STANDBY UNTIL YOUR JOB HAS BEEN SUBMITTED. RELAX.  THIS MAY TAKE A MINUTE.
PREPARING OUTPUT JCL AT THIS POINT.  DON'T LEAVE NOW.
JOB ERS7XX01(JOB03527) SUBMITTED  **  FREE ALL FILES  **
YOUR JOB HAS BEEN HAS BEEN SUBMITTED AS A CLASS D JOB FOR YOU ERS188
IF YOU FEEL THAT YOU HAVE MADE A MISTAKE ENTER:
CANCEL ERS7XX01  __ AND START OVER AGAIN.
TO CHECK THE STATUS OF YOUR JOB, ENTER:
ST ERS7XX01
ENTRY (A) ERS188.ER7XX.DUMMY.SPSS.SUBMIT.CNTL DELETED
READY
```

JOB CONTROL LANGUAGE OUTPUT FROM EXAMPLE 4

```
1      //ERS7CG01   JOB (                ,RJO13),'SPSS RUN',CLASS=D,
       // TIME=(2,0),PRTY=3,MSGCLASS=Q,NOTIFY=ERS34
       ***ROUTE PRINT RMT13
2      //STEP1   EXEC PGM=SPSS,REGION=300K,PARM=160K
3      //STEPLIB DD DSN=SYS1.SPSSH,DISP=SHR
4      //FT10F001 DD DSN=SYS3.SPSSUDCU.DATA,DISP=SHR,DCB=BUFNO=1
5      //FT01F001 DD UNIT=SYSDA,SPACE=(CYL,(2,1))
6      //FT02F001 DD UNIT=SYSDA,DCB=BLKSIZE=2012,
       //   SPACE=(CYL,(5,1))
7      //FT05F001 DD UNIT=SYSDA,DISP=SHR,
       //   DSN=ERS99.ER7XX.MERGE.FILES.CONTROL.DATA
8      //FT06F001 DD SYSOUT=T,DCB=(RECFM=VBA,LRECL=137,BLKSIZE=826)
9      //FT03F001 DD UNIT=SYSDA,DISP=SHR,
       //   DSN=ERS99.ER7IN.SYSTEM.FILE1.DATA,
       //   LABEL=(,SL,,IN),
       //   VOL=SER=
10     //FT03F002 DD UNIT=SYSDA,DISP=SHR,
       //   DSN=ERS99.ER7IN.SYSTEM.FILE2.DATA,
       //   LABEL=(,SL,,IN),
       //   VOL=SER=
11     //FT04F001 DD UNIT=SYSDA,DISP=(NEW,CATLG),
       //    DSN=ERS99.ER7OT.MERGED.SYSTEM.FILE.DATA,
       //    SPACE=(TRK,(11,5),RLSE)
12     //FT09F001 DD UNIT=TAPE,DISP=(NEW,KEEP),
       // DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400),
       //   LABEL=(,SL,,,RETPD=0060),
       // DSN=NULLFILE
13     //FT11F001 DD UNIT=SYSDA,SPACE=(CYL,(10,10)),
       //   DSN=&TEMP
14     //FT12F001 DD UNIT=SYSDA,SPACE=(CYL,(5,10))
15     //SORTLIB   DD DSNAME=SYS1.SORTLIB,DISP=SHR
16     //SYSOUT    DD SYSOUT=T
17     //SORTWK01 DD UNIT=SYSDA,SPACE=(TRK,(50),,CONTIG)
18     //SORTWK02 DD UNIT=SYSDA,SPACE=(TRK,(50),,CONTIG)
19     //SORTWK03 DD UNIT=SYSDA,SPACE=(TRK,(50),,CONTIG)
20     //SORTWK04 DD UNIT=SYSDA,SPACE=(TRK,(50),,CONTIG)
```
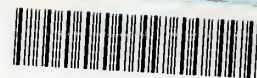
IV. SUMMARY

SPSS is highly useful to the Economic Development Division. Its ease of use, as well as the extensive statistical procedures, make it the most heavily utilized statistical package by EDD analysts. The research staff continue to be loyal users and look forward to enhancements in future releases.